Published in final edited form as:

Nat Methods. 2013 June ; 10(6): 453–454. doi:10.1038/nmeth.2453.

Comprehensive objective maps of macromolecular conformations by quantitative SAXS analysis

Greg L. Hura¹, Helen Budworth², Kevin N. Dyer¹, Robert P. Rambo¹, Michal Hammel¹, Cynthia T. McMurray², and John A. Tainer^{2,3}

¹Physical Bioscience Division, Lawrence Berkeley National Lab, 1 Cyclotron Rd. Berkeley CA 94720

²Life Science Division, Lawrence Berkeley National Lab, 1 Cyclotron Rd. Berkeley CA 94720

³Department of Molecular Biology and The Skaggs Institute of Chemical Biology, The Scripps Research Institute, La Jolla, CA

Abstract

Comprehensive perspectives of macromolecular conformations are required to connect structure to biology. Here we present a small angle X-ray scattering (SAXS) Structural Similarity Map (SSM) and Volatility of Ratio (V_R) metric providing comprehensive, quantitative and objective (superposition-independent) perspectives on solution state conformations. We validate V_R and SSM utility on human MutS β , a key ABC ATPase and chemotherapeutic target, by revealing MutS β DNA sculpting and identifying multiple conformational states for biological activity.

Biological macromolecules function by adopting conformationally-distinct states through processes such as phosphorylation, nucleotide binding, ATP hydrolysis, ligand binding, complex formation, or combinatorial post-translational modification. For example, the Rad50 ABC ATPase complex with Mre11 and Nbs1 is a trimer of dimers: 6 subunits each with at least 3 distinct states for a possible total of 216 states¹. Many of these conformational states may be biologically important, but relatively few are observed and compared by current structural techniques. Crystallography is precise but low throughput, typically captures one low-energy conformation, and often requires truncations or mutations. In contrast, SAXS experiments can 1) probe the solution state under any condition, 2) provide information at resolutions sufficient to distinguish conformational states², 3) characterize flexible macromolecules and 4) screen in high throughput³. Unfortunately, current analytical methods have been low throughput and under-developed, typically requiring a visual one-by-one comparison for characterizing possible differences between solution states. The need for global conformational comparisons prompted us to develop a structural similarity map (SSM) for rapid SAXS screening. SSM is an analytical and visual tool that both discriminates “at a glance” and quantifies the conformational similarities and differences among large conformational arrays.

An SSM is a $n \times n$ diagonally symmetric matrix, where n is the number of SAXS data sets collected, and each matrix cell quantitates pairwise agreement between SAXS curves. The similarity score is displayed as a gradient color (white being low and red being high). SSM

Correspondence to: Cynthia T. McMurray; John A. Tainer.

Additional material. A supplementary methods document contains supporting figures, detailed biochemistry and analytical methods.

Author Contributions G.L.H., J.A.T., M.H., R.P.R. and K.N.D. developed SSMs and V_R . H.B. and C.T.M. prepared samples for data collection and aided in writing manuscript.

utility is enhanced by our new Volatility of Ratio (V_R) metric. Importantly, V_R can discriminate local changes acting in biological functions where most of the structure remains the same. The V_R metric is calculated by taking the ratio of two SAXS profiles, normalizing the ratio so the average over the range is 1 and binning the resulting ratio at a minimal frequency⁴ ($\Delta q = \pi/d$ where q is a function of scattering angle θ , and X-ray wavelength λ , $q = 4\pi(\sin(\theta/2))/\lambda$). By assuming a maximum dimension $d < 40\text{nm}$ the number of bins N is 25 over a q range $q < 0.2\text{\AA}^{-1}$. The volatility of the binned ratio is then calculated by taking the sum of the absolute value of the difference between sequential bins divided by their

average (Supplemental Fig. 1): $V_R = \sum_{i=1}^{14} \text{abs} \left[\frac{R(q_i) - R(q_{i+1})}{(R(q_i) + R(q_{i+1}))/2} \right]$ where R is the ratio of the intensities at q_i .

Properties of X-ray scattering inspired V_R . The Debye approximation, describing elastic X-ray scattering, is modular as used for form factors $f(q)$. If much of the structure stays the same (domains, folds, subunits), one may utilize $f(q)$ s that encompass all atoms within these structural units. By dividing one curve by another, the dominating and exponentially decaying $f(q)$ contributions are largely removed: all of q space becomes equally weighted. Changes in structure are apparent in SAXS curves as differences in oscillations about the exponential decay. Differences often remain subtle, so binning reduces noise in each point through averaging while providing sufficient discrimination by our measure of volatility.

Utilizing V_R for scoring and a color gradient for display effectively and rapidly discriminates structural and conformational similarity among populations of macromolecules under different solution conditions. The “vector components” of the SSM scores the agreement of a single reference experiment against others (Fig 1, Supplemental Fig. 2). χ^2 is often used to provide binary “accept or reject” criteria for judging whether a proposed atomic model is consistent with experimental SAXS data⁵. Alternate metrics identify⁶ or predict⁷ atomic resolution structures by comparing calculating profiles against experiment. We optimized V_R not only to distinguish when structures are the same or different, but also to show the degree of similarity. SAXS is a probe of electron density pair correlations both at global and local scales, and objectively characterizes similarity, more so than many commonly applied metrics. For example, the SSM in Figure 1a compares structural snapshots from a molecular dynamics simulation following phosphorylation dependent decoupling of retinoblastoma protein domains⁸. SAXS metrics show greater agreement between snapshots taken closer in time; following the simulation. In contrast, a root means square deviation (RMSD) comparison does not correlate with the simulation. RMSD requires a superposition, which biases assessment of similarity. SAXS is a superposition-independent averaging technique, so comparisons are unbiased and can also detect conformational shifts in a small population (Fig 1b).

V_R identifies structural similarity under conditions inducing multimerization or up to 25% truncations (Fig 1c, Supplemental Fig. 3). Advantages of V_R over χ^2 can be subtle, e.g. both distinguish tense and relaxed states of aspartate transcarbamoylase conformations (Fig 1b). However, optimally, V_R highlights key distinctions that are missed by the other metrics (asterisk in Fig 1c). Notably, V_R identifies a multimer and a truncation as more similar to the reference structure relative to a protein of similar size but of a different fold.

SSMs provide a comprehensive view of high-throughput SAXS data collected from multi-component biological systems. SSMs are built from the vectorial components, described above. We tested and validated SSM utility for global perspectives of conformations on experimental data from a key DNA repair ABC ATPase MutS β . MutS β removes extrahelical small DNA loops and improperly paired DNA bases⁹, which, if unrepaired, lead to cancer and genome instability. Human MutS β (Fig. 2a) is a heterodimer comprising one

subunit each of MSH2 (104 kDa) and MSH3 (127 kDa), both of which undergo an ordered series of nucleotide-dependent steps to initiate repair. Each discrete nucleotide-bound state is a conformational decision point that “sets-up” coupling to the next step along the repair pathway, and must be preserved. Thus, a mechanistic understanding of these steps is key for elucidating how cells avoid mutation.

We measured 10 liganded states of MutS β with and without DNA substrate and in the presence of four distinct nucleotides or ATP analogues (Figure 2b, Supplemental Fig. 4 and 5). Since DNA lesion recognition depends on ATP hydrolysis, non-hydrolysable analogues and ADP were used to probe conformational states before and after hydrolysis, respectively. Three features in the SSM are apparent (Fig. 2b). First, the color patterns of rows fall into three categories indicating three main functional nucleotide states. Second, matrix cells comparing ADP and ATP conformations are scored as similar and contrast with non-hydrolyzable ATP analogues. This is true both with and without DNA (regions 1 and 2 in Fig. 2b) implying MutS β rapidly hydrolyzes ATP to ADP. Crosslinking of radiolabeled nucleotides verified that active ATP hydrolysis in MSH3 converted ATP to ADP (Fig. 2c, γ ATP Mg²⁺), and that both subunits were occupied whether ADP or ATP was the added nucleotide (Fig. 2d, α -ATP and α -ADP). Third, DNA has surprisingly little impact on MutS β conformational status (asterisk in region 3 in Fig 2b). MutS β remained DNA-bound independent of the added nucleotide (Fig. 2c). These results imply that DNA is sculpted to the protein conformation, and nucleotide binding is sufficient to drive the MutS β conformational changes. Indeed, in smFRET¹⁰ and crystal structure¹¹, MutS β binding bends DNA, but the MutS β conformation is similar on and off DNA. That MutS β sculpts the DNA to the protein conformation has implications for mismatch repair and signaling, as DNA sculpting can control pathway progression for base and nucleotide repair¹².

SSMs can guide further investigation, either by identifying contrasting states or by showing trends. For example, further investigation of the SAXS profiles from the contrasting states (Fig. 2) with ATP vs AMPPNP show a 6 Å difference in radius of gyration (R_g). The AMPPNP inhibited transition state complex is more compact. The need for direction provided by SSMs increases as data from high-throughput screening of other DNA substrates, potential chemotherapeutic inhibitors, MutS β patient mutations, and its homologue MutS α are added. By identifying similar states through clustering, biologically important distinctions in subtle conformational changes and trends may be revealed.

Despite enormous advances in structural biology, there remains a desperate need for methods allowing a more comprehensive characterization of proteins including their biologically relevant complexes and conformations¹³. SSM is an objective, global view of large SAXS datasets defining conformations for a macromolecule in a host of contexts. The SSM not only provides insight into the full range of macromolecular conformations, but offers discovery potential for conformational aspects not observed by other methods. For example, the independence of MutS β conformations on DNA was not evident from biochemical measurements but obvious by SSM. We anticipate wide spread use of these maps as the analysis is robust and SAXS data collection is increasingly accessible. We, therefore, developed a web application (http://sibyls.als.lbl.gov/saxs_similarity) for creating SAXS SSMs currently displaying up to 24 scattering profiles. Each individual cell may also be selected to report on the change in the radius of gyration (ΔR_g) of any two profiles. We expect to add criteria such as changes in mass, volume, surface area and flexibility, as the web application develops. We have also made available open source code for creating matrices involving hundreds of SAXS profiles based on V_R or Chi^2 . Looking ahead, planned next generation light sources for SAXS measurements will provide data from sub-microliter volumes of sub-microgram quantities on sub-second data collection times. Thus, biological sample size will decrease sharply while throughput will increase. Capitalizing on these

capabilities requires high-throughput and comprehensive structure evaluation methods. With the discriminating power of V_R , SSMs objectively assess similarity among macromolecular states unbiased by superposition, aids visualization and provides global perspectives on functional conformations connecting structures to biological outcomes.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Support for advancement of SAXS technologies at the Lawrence Berkeley National Laboratory SIBYLS beamline of the ALS came from the DOE program Integrated Diffraction Analysis Technologies (IDAT) and Laboratory Directed Research and Development (LDRD) under Contract DE-AC02-05CH11231 with the U.S. Department of Energy.

References

1. Williams GJ, Lees-Miller SP, Tainer JA. Mre11-Rad50-Nbs1 conformations and the control of sensing, signaling, and effector responses at DNA double-strand breaks. *DNA Repair*. 2010; 9:1299–1306. [PubMed: 21035407]
2. Putnam CD, Hammel M, Hura GL, Tainer JA. X-ray solution scattering (SAXS) combined with crystallography and computation: defining accurate macromolecular structures, conformations and assemblies in solution. *Quarterly Reviews of Biophysics*. 2007; 40:191–285. [PubMed: 18078545]
3. Hura GL, et al. Robust, high-throughput solution structural analyses by small angle X-ray scattering (SAXS). *Nature Methods*. 2009; 6:606–U83. [PubMed: 19620974]
4. Moore PB. Small-Angle Scattering - Information-Content and Error Analysis. *Journal of Applied Crystallography*. 1980; 13:168–175.
5. Svergun D, Barberato C, Koch MHJ. CRY SOL - A program to evaluate x-ray solution scattering of biological macromolecules from atomic coordinates. *Journal of Applied Crystallography*. 1995; 28:768–773.
6. Sokolova AV, Volkov VV, Svergun DI. Database for rapid protein classification based on small-angle X-ray scattering data. *Crystallography Reports*. 2003; 48:959–965.
7. Yang YD, Faraggi E, Zhao HY, Zhou YQ. Improving protein fold recognition and template-based modeling by employing probabilistic-based matching between predicted one-dimensional structural properties of query and corresponding native properties of templates. *Bioinformatics*. 2011; 27:2076–2082. [PubMed: 21666270]
8. Balog ERM, Burke JR, Hura GL, Rubin SM. Crystal structure of the unliganded retinoblastoma protein pocket domain. *Proteins-Structure Function and Bioinformatics*. 2011; 79:2010–2014.
9. Owen BAL, Lang WH, McMurray CT. The nucleotide binding dynamics of human MSH2-MSH3 are lesion dependent. *Nature Structural & Molecular Biology*. 2009; 16:550–557.
10. Lang WH, et al. Conformational trapping of Mismatch Recognition Complex MSH2/MSH3 on repair-resistant DNA loops. *Proceedings of the National Academy of Sciences of the United States of America*. 2011; 108:E837–E844. [PubMed: 21960445]
11. Gupta S, Gellert M, Yang W. Mechanism of mismatch recognition revealed by human MutS beta bound to unpaired DNA loops. *Nature Structural & Molecular Biology*. 2012; 19:72–U91.
12. Tubbs JL, et al. Flipping of alkylated DNA damage bridges base and nucleotide excision repair. *Nature*. 2009; 459:808–813. [PubMed: 19516334]
13. Harrison SC. Comments on the NIGMS PSI. *Structure*. 2007; 15:1344–1346. [PubMed: 17997957]
14. Jin L, Stec B, Lipscomb WN, Kantrowitz ER. Insights into the mechanisms of catalysis and heterotropic regulation of Escherichia coli aspartate transcarbamoylase based upon a structure of the enzyme complexed with the bisubstrate analogue N-phosphonacetyl-L-aspartate at 2.1 angstrom. *Proteins-Structure Function and Genetics*. 1999; 37:729–742.

Figure 1.

SAXS as a measure of structural similarity; scored by three metrics (χ^2 , Pearson⁷, and V_R). Scores were assigned a gradient color with white - low agreement and red - high agreement. Right-most box in each panel are self-comparisons. (a) Phosphorylation state modulates the coupling of domains in Retinoblastoma protein⁸. Snapshots from a molecular dynamics simulation of this process were taken (right to left) 1000, 2000, 4000, 8000 and 16000 time steps away from the reference structure. The closer in time the snapshots are, the more correlated their structure. (b) Aspartate Transcarbamoylase (ATCase) is a canonical cooperative enzyme¹⁴ with available high resolution structures of a tense state (T), relaxed state (R) and mutant intermediate. Calculated SAXS curves from the three structures and mixtures of T and R states were judged for agreement against reference structure T-state (c) Comparing a protein structure against (right to left) another with identical number of amino acids but different fold, their dimers, and 25% truncations. *indicates where the trends in metrics disagree.

Figure 2.

A SAXS structural similarity map (SSM) for MutSβ with and without ADP, ATP, ATP analogs and DNA. A schematic of the complex is shown in a). The SAXS SSM is shown in b). The diagonal (black squares) compares a SAXS curve to itself. * compares states with vs without DNA c) Non-denaturing gel electrophoresis of the samples show DNA binding. (d) Ultraviolet crosslinking following ATP hydrolysis.